DOCUMENT RESUME

ED 069 730 TM 002 182

AUTHOR	Freeberg, Norman E.; Reilly, Richard G.
TITLE	Validation of a Test Battery for Youth-Work Training Program Enrollees. Research Memorandum.
INSTITUTION	Educational Testing Service, Princeton, N.J.
SPONS AGENCY	Greater Chester Movement, Chester, Pa.; Mercy-Douglas Hospital, Philadelphia, Pa. Neighborhood Youth Corps Project.
REPORT NO	ETS-RM-72-11
PUB DATE	Aug 72
NOTE	28p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Adolescents; Attitude Tests; Cognitive Tests; Correlation; *Criterion Referenced Tests; Data Analysis; Disadvantaged Youth; Dropout Programs; Negro Youth; Performance Tests; *Predictive Ability (Testing); *Questionnaires; Research Methodology; Tests; *Test Validity; Trainees; Work Experience Programs
IDENTIFIERS	*Enrollee Test Battery

ABSTRACT

Correlations between the Enrollee Test Battery (ETB) and performance on more conventional tests were examined to determine the ETB validity, in addition to a small-Scale study of predictive validity of both ETB and the conventional tests. Three types of paper and pencil measures were utilized: the ETB, conventional cognitive skills tests, and scales of criterion questionnaires. There were 23 tests in all. The study sample consisted of 74 (29 males and 45 females) predominantly black high school dropouts, between 14 and 17 years of age, who were enrolled in two Neighborhood Youth Corps Out-of-School projects. Answers to the criterion questionnaires were obtained from 44 of the enrollees six months after they had left the program. Predictive validity for the tests was examined by correlating each of the 23 tests scores with factor scores derived from the questionnaire criterion scales for each of the three criterion samples--Program Completion, Post Program (Employed), and Post Program (Not Employed). Levels of concurrent validity for the ETB was determined from intercorrelations between the 17 tests of that battery and the 6 conventional cognitive skill measures. That matrix was also factor analyzed. Results of the study show that measures of the ETB, which was designed specifically for use with disadvantaged adolescents, appear fairly coherent and logical in their patterns of relationships with each other and with conventional cognitive skills tests. (DB)

ED 069730

RM-72-11

RESEARCH MEMORANDUM

TI 002 182

VALIDATION OF A TEST BATTERY FOR YOUTH-WORK TRAINING PROGRAM ENROLLEES

Norman E. Freeberg

and

Richard R. Reilly

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

This Memorandum is for interoffice use. It is not to be cited as a published report without the specific permission of the authors.

Educational Testing Service
Princeton, New Jersey
August 1972

Validation of a Test Battery for Youth-Work Training Program Enrollees[1]

Norman E. Freeberg

and

Richard R. Reilly

Background and Purpose

A battery of paper and pencil measures, designed for use as guidance tools with youth-work training program enrollees, had undergone initial development on the basis of data obtained from adolescent male and female school dropouts in the Neighborhood Youth Corps Program (Freeberg, 1970). Measures comprising that experimental battery (Enrollee Test Battery) were tailored, in terms of content, verbal level, format and method of administration, to overcome a number of deficiencies often attributed to conventional formal tests when applied to youngsters from poverty-level backgrounds who are also minority group members with low verbal skill levels (Eells et al., 1951; Karp & Sigel, 1965; SPSSI, 1964). The particular measurement constructs chosen for incorporation in the Enrollee Test Battery were intended to reflect trainee behaviors considered relevant to work training program purposes and to the informational needs of program professionals (e.g., guidance counselors). Those constructs are embodied in a variety of attitudinal, practical reasoning and vocational orientation measures found in 13 separate test booklets that combine pictorial and verbal information. The booklets are relatively short

---

[1]Funds for this study were provided in part by the Neighborhood Youth Corps projects of Mercy-Douglas Hospital, Philadelphia, Pa., and the Greater Chester Movement of Chester, Pa. The authors acknowledge the invaluable support of the directors of those N.Y.C. projects, Mr. Eben Short and Mr. Norman Watts, as well as the assistance of the project counselors.

in length (i.e., 10 to 28 items), are presented without time limits and require oral presentation by the examiner.

In the initial phase of development, the suitability of the measures was determined by analyses of their item characteristics, estimates of internal consistency (reliability), the pattern of relationships between the measures of the battery (factor structure) and the validity of each measure against concurrently obtained criterion rating scores from program guidance counselors and work-site supervisors. Although a number of the tests showed reasonable promise on the basis of their measurement properties, broader determinations of validity, both concurrent and predictive, are clearly in order before the measures could prove professionally defensible for application to disadvantaged adolescent groups.

The present study is undertaken as a next-step in what should be a continuing process of meeting such validation needs by demonstrating test relationships to a variety of relevant performance criteria. A primary intention is to examine--for the first time--correlations between the newly-devised Enrollee Test Battery and performance on more conventional, formal tests, some of which have either been used extensively with disadvantaged young adults or carry explicit assertions, by the publisher, of their applicability to such a population. The set of conventional measures chosen for that purpose is intended to span a range from verbal and conceptual abilities (e.g., vocabulary, arithmetic, reading comprehension) to nonverbal reasoning (e.g., figure analogies) and associative, rote-memory skill.

In addition to validation based on concurrent test performance, there is a limited opportunity to examine predictive validity for the experimental test battery as well as for the published measures. This is to be done by

utilizing criterion data obtained for relatively small samples of youth work training program enrollees at the time they completed their stay in the program and again some six months after they had left.

Because of expected restrictions in sample size, geographic locale and the scope of criterion information available, this predictive validity phase should be viewed solely as a small-scale, or pilot-study effort. It is attempted here in order to provide preliminary results, as well as needed experience with the difficult follow-up and data gathering procedures required for a sample of minority group, school dropouts residing in inner-city ghetto areas. Such information is expected to be of value in planning further larger-scale validation studies.

Method

Description of the Measures

Described below are the essential features of the three types of paper and pencil measures utilized: the Enrollee Test Battery, the conventional cognitive skill tests and the scales of the criterion questionnaires.

A. Enrollee Test Battery (ETB). Only a brief description of the ETB, designed explicitly for disadvantaged adolescents, is either warranted or practical for purposes of this paper. More complete descriptions of the rationale, design and measurement characteristics are available elsewhere (Freeberg, 1968, 1970).

Of 13 measures originally developed, 11 were considered acceptable, by virtue of content and psychometric characteristics, for this validation study. Two are measures of practical reasoning skill:

(1) Practical Reasoning - Zip Coding (10 items): provides the respondent with information needed for a hypothetical post office job requiring him to sort mail using zip codes. Questions are presented in a multiple-choice format pertaining to that task.

(2) Practical Reasoning - Map Reading (10 items): presents a map illustrating several square city blocks along with information needed to answer multiple choice questions about delivering materials to different locations shown on the map.

Measures of vocational orientation consist of:

3) Job Knowledge (27 items): requiring answers to multiple-choice items regarding a variety of jobs (e.g., carpenter, auto mechanic, policeman) in terms of their educational requirements, starting salary, primary task performed, hours of work, place where work is performed and tools utilized.

(4) Job Seeking Skills (17 items): presents multiple-choice items dealing with ways of looking for jobs, some of which entail interpretation of newspaper want ads and portions of job application blanks.

(5) Job Holding Skills (11 items): which depicts situations that might be encountered on a job with regard to supervisors' requests, appropriate dress, punctuality, etc., for which the respondent chooses the reply he would give, or the most appropriate behavior called for, in that situation.

(6) Vocational Aspirations minus Plans (VA-VP Discrepancy) (16 items): is derived from two test booklets, both showing the same jobs (pictorially and by designation). The respondent indicates

the degree to which he would aspire (prefer) to engage in any particular occupation and then the degree to which he believes it is likely that he would enter that same occupation. The discrepancy between scores for aspiration and plans was utilized for all analyses.

Attitudinal measures of the battery deal with:

(7) Attitude Toward Authority (12 items): in which scenes depicting responses of an adolescent to authority figures (e.g., teacher, policeman, parents, judge, etc.) are used to elicit the degree of pro- or anti-authority feelings by the respondent.

(8) Self-Esteem (16 items): depicts situations pictorially in which the respondent indicates the degree to which he feels himself "worthy" of the desirable outcomes (getting a diploma, job, promotion, etc.).

(9) Motivation for Vocational Achievement (10 items): presents statements (item stems) that bear on the desirability of seeking and maintaining employment. These require the individual to respond (Likert-Scale format) in terms of the degree of agreement with the statement.

(10) Deferred Gratification (16 items): presents statements dealing with the willingness to delay present reward for future gain to which the respondent indicates his degree of agreement.

A final test booklet in the battery covers:

(11) Vocational Interest (28 items): which presents pictorial and verbal information regarding job tasks characteristic of a number of occupations. The respondent indicates the degree to which he

could favor performing each task. Separate scale scores are derived for each of seven interest categories (4 items each) designated as Clerical, Service, Technical, Outdoor, Science, Business and Aesthetic.

B. Formal cognitive skills measures. In addition to the above Enrollee Test Battery, six measures of the more conventional, published variety of paper and pencil tests served as the concurrent criteria. One is the Paragraph Meaning subtest of the Stanford Achievement Test, Form W, Intermediate I (Harcourt, Brace, & World, 1964), which has been used extensively for determining reading levels of youngsters from all socioeconomic levels and in a variety of educational settings. Two measures, designated as the Arithmetic and Vocabulary subtests, were selected from the Adult Basic Learning Exam (ABLE), Level II, Form A (Harcourt, Brace, Jovanovich, 1970). These ABLE measures appear largely conventional in content and format although they are said to be designed, according to the publishers, for adolescents from educationally disadvantaged backgrounds (Karlsen, Madden, & Gardner, 1971).

The three remaining measures are all nonverbal subtests drawn from a battery with known factorial properties (French, Ekstrom, & Price, 1969). These consist of a conventional test of Figure Analogies, a maze-following type of measure designated as Choose-A-Path and an associative (rote) memory test designated as Picture-Number. The three measures have also been utilized previously, as part of a guidance and research program with junior college populations (College Entrance Examination Board, 1969) and with 11th and 12th grade inner-city students (Flaugher, 1971).

C. Criterion measures. For purposes of determining the predictive validity of all the above paper and pencil measures, scores were obtained from questionnaires developed in a previous study of criteria for youth-work training programs (Freeberg & Reilly, 1971). Based on factors obtained in that study, sets of variables were scored to yield: (a) three coherent and relatively independent criterion scales for use at program completion, (b) another set of three at post-program periods for those former enrollees who had obtained employment, and (c) a third set of three criterion scales for former enrollees who had not obtained employment over the defined post-program period.

.Program Completion Criterion Scales are designated as: (1) "Training Program Adjustment"--which combines variables of trouble with police, counselor proficiency ratings, work supervisor proficiency ratings, worksite absences, number of jobs in which the enrollee is willing to take training; (2) "Social Adjustment"--combining variables of problems with people in the community, family adjustment, number of police contacts, peer rating score; and (3) "Job Aspiration Level"--based on variables of enrollee salary expectations, level of a highly desired job, level of long-range plans.

.Post-Program Criterion Scales consist of one set of scales for those who had obtained full-time employment (over a period of one week or more) after leaving the program. These are defined as: (1) "General Job Success and Adjustment"--using variables of employer proficiency rating, financial savings, whether or not employed at the time of the interview, length of stay on job, family adjustment, police contacts; (2) "Striving for Vocational Success"--which combines variables of job promotion, amount of first raise, problems with

people in the community, quality of next job desired, ease with which contacted by the interviewer; and (3) "Stability-Mobility"--based on combining variables of number of places interviewed for jobs, number of jobs held, number of visits to State Employment and number of places lived.

Applicable post-program scales, for those who had not obtained employment, were derived from relatively weak criterion factors that were found to be neither as coherent nor as interpretable as the sets of factors for the other groups (Freeberg & Reilly, 1971). These are: (1) "Community Adjustment"--which combines variables of trouble with police, willingness to contribute financially to family, ease of obtaining credit; (2) "Striving for Personal Improvement"--with variables of part-time work experience, salary expectations, ease with which contacted by interviewer; and (3) "Realistic Aspirations"--combining variables of estimated ability to perform on a sought-after job, knowledge of the job sought, salary expectations, and family adjustment.

Sample and Data Collection

Twenty-nine males and 45 females enrolled in two Neighborhood Youth Corps Out-of-School projects, located in a northeastern urban area, constituted the sample from which the test data were obtained. The sample of 74 were predominantly Black high school dropouts, between 14 and 17 years of age, who had been enrolled for approximately several weeks, or less, at the time they were tested by the project guidance counselors. Of the 74, a group of 20 who remained enrolled for approximately 5 to 6 months or more (arbitrarily defined as program "completion") could be contacted at the time they left the program and were also willing to respond to an individually administered program-completion criterion questionnaire.

Forty-four of the 74 enrollees were located approximately six months after they had left the program and consented to respond to the individually administered post-program criterion questionnaire. Seventeen of these 44 former enrollees obtained full-time employment (Post-Program: Employed Sample) and 27 had failed to obtain full-time jobs lasting for a period of one week or more (Post-Program: Not Employed Sample). Enrollees were paid for responding to the criterion questionnaires.

Data Analyses

Levels of concurrent validity for the Enrollee Test Battery were determined from intercorrelations between the 17 tests of that battery and the six conventional cognitive skill measures, as obtained from a 23 x 23 intercorrelation matrix of all test scores. That matrix was also factor analyzed, as a way of summarizing overall patterns of relationships, using a principal axis solution with Varimax rotation of the factors to orthogonality (Kaiser, 1958).

Predictive validity for the 23 tests was examined by correlating each of the 23 test scores with factor scores derived from the applicable questionnaire criterion scales for each of the three criterion samples--i.e., Program Completion, Post-Program (Employed), and Post-Program (Not Employed). Each factor score consisted of a weighted sum of the standardized variables with loadings greater than .30. In all cases the weights used were the factor loadings derived in a previous analysis of the questionnaire with a larger, more representative sample (Freeberg & Reilly, 1971) than the present one.

Since missing data occurred throughout the samples for the test scores (i.e., N's ranged from 67 to 74), computation of the intercorrelation matrix

required the use of missing data analyses. For the much smaller available samples, used to determine predictive validities, missing data resulted in varied sample sizes (N's ranged from 10 to 27). Where differing N's occur, they are appropriately indicated in the tabled results.

Results and Discussion

Concurrent Validation

The matrix showing the intercorrelations for the 23 tests is presented in Table 1.

Insert Table 1 about here

Moderate to substantial levels of concurrent validity occur most consistently between four subtests of the ETB and five of the six conventional measures of cognitive skill. These are found, appropriately enough, for the four subtests of the ETB that possess an obvious requirement for reasoning or verbal-intellectual capabilities (i.e., Job Knowledge, two Practical Reasoning subtests and the Job Seeking Skills test). Within that set of coefficients--bounded by heavy lines in the upper-right portion of the matrix--it can be seen that the Job Seeking Skills measure results in the highest levels of concurrent validity overall, especially with the more complex conceptual skills involving word meaning ($r = .57$), arithmetic ($r = .58$) and paragraph comprehension ($r = .75$).

It should be noted that, in general, the levels of intercorrelation among the six conventional measures (variables #18 to #23) are no higher than their correlations with the four cognitive skill measures of the ETB; although

the latter are considerably different in design and comparatively much shorter in length.

One striking feature of the relationships between the cognitive skill measures is seen in the correlations of low magnitude between the Picture-Number Test and the other five conventional measures, as well as with the four cognitive skills subtests of the ETB (r's ranging from .05 to .33 with only three of the nine correlations reaching significance). Those uniformly low correlations for this sample of minority-group adolescents reflect a degree of relative independence for the purely associative, rote-memory task. The finding is somewhat consonant with differences in learning strategies that have been suggested for the acquisition of rote learning, in contrast to strategies believed to underlie higher-level conceptual abilities (Rohwer, 1971). Moreover, when rote-memory test scores are correlated with measures of conceptual skills for different ethnic groups, the lowest of the correlations (i.e., the greater relative uniqueness of rote memory skill) have been found for Black adolescents from inner-city schools (Flaugher & Rock, 1972).

Significant correlations between the attitudinal scales of the battery and the cognitive-intellectual measures (Table 1) are relatively few and scattered with no clear pattern evident. Only the several positive r's in the .20's and .30's appearing among the Job Holding Skills, Attitude Toward Authority and Motivation for Vocational Achievement scales suggest possible consistencies of these attitudes in relation to cognitive skills. For the most part, however, those results approximate the extent of relationships expected when attitudinal and cognitive-intellectual measures are inter-correlated.

A much broader perspective of the correlational patterns is provided by a factor analysis of the matrix of Table 1, which also serves as a form of construct validation for all of the measures used. Loadings on the four factors extracted by that analysis (Table 2) confirm, in a more coherent way, much of what has already been discussed on the basis of visual examination of the individual correlations.

Insert Table 2 about here

Thus, Factor I, Cognitive Ability, emerges as one that is clearly dominated by high loadings on all of the cognitive skills measures with the exception of the Picture-Number Test, which shows relatively little presence on that dimension. Of the loadings for the noncognitive (attitude or interest) scales only the Attitude Toward Authority scale reaches an interpretable level (i.e., .30), suggesting that there is a tendency for those who express greater acceptance of conventional authority to achieve better intellectual skills scores.

Factor II represents an unquestioned dimension of generalized Interest in Vocational Tasks--of all sorts. Coherence of this factor is consistent with the findings for similar job-task interest scales using a wider range of ethnic groups and SES levels in the samples tested (Knapp, Grant, & Demos, 1966). Equally consistent with previous findings is the relative independence of the interest factor in relation to attitudinal and intellectual skills measures (Cronbach, 1960; Tyler, 1956).

The attitudinal scales that comprise Factor III are reasonably interpretable as a dimension of Positive Social Attitudes, with an apparent

vocational focus. Thus, those enrollees who possess stronger feelings of Self-Esteem are the ones who also score higher on the Motivation for Vocational Achievement scale, indicate a higher degree of awareness of proper on-the-job behaviors (Job Holding Skills) and tend to prefer Service job tasks found in occupations that entail dealing with, or helping, others (e.g., social worker, store clerk). These attitudinal scales are relatively independent of any intellectual skills measures with the exception of the positive loading found for the Arithmetic Test (.37).

The last factor extracted, Factor IV, is interpreted most sensibly as one that differentiates Masculine Vs. Feminine Interests and is a likely result of having combined males and females into a single sample. Aspects of the loading pattern that support the interpretation stem from the dominant positive loadings on Technical and Outdoor Vocational task interest--known to result in higher scores for males (Freeberg, 1968)--accompanied by negative loadings on the more female-oriented Clerical job tasks. At the same time, the factor depicts an enrollee who performs more poorly in paragraph comprehension and also shows less willingness to defer present gratification for future gain.[2] There is also a tendency for the individual who is on the masculine-interests side of this dimension to perform better on the nonverbal Choose-A-Path test.

---

[2]Female school dropouts from urban areas who enter youth-work training programs are generally found to be superior to males in reading ability and also tend to score higher on this particular Deferred Gratification Scale (Freeberg, 1968).

Predictive Validation

Predictive validities of the 23 test scores are presented in Tables 3, 4, and 5 for the Program Completion, Post-Program (Employed), and Post-Program (Not Employed) criterion samples respectively.

Insert Tables 3, 4, and 5 about here

In what must necessarily be a very cautious and broad interpretation of these small-sample results, it is apparent that the most consistent patterns and highest levels of significant validity are dominated by the cognitive-intellectual tests. Few scattered validities for the interest or attitudinal scales appear across the three criterion samples, although those few significant ones serve as an aid in interpreting the overall findings. Specifically, among the validities for the Program Completion criteria (Table 3), "Social Adjustment" (with regard to family, peers, police, other members of the community) is best predicted from the higher level conceptual skills of Arithmetic ($r = .66$), Paragraph Meaning ($r = .52$), and Practical Reasoning: Zip Coding ($r = .51$). That criterion dimension is also associated positively with Self-Esteem scores ($r = .47$) and negatively ($r = -.44$) with the level of interest shown for Outdoor Job Tasks (i.e., lower level, farm-related tasks requiring physical labor).

The other set of validities with some degree of consistency in their pattern of significance is, again, found to occur for a number of cognitive skills tests. These appear for the Post-Program (Employed) enrollee sample as predictive of the "Job Success and Adjustment" criterion scale (Table 4). In this instance, significant r's occur over a fairly wide range of cognitive skills encompassing the purely nonverbal Figure Analogies ($r = .57$) and

Picture Number Tests (r = .51), as well as the more verbally demanding Vocabulary (r = .49), Arithmetic (r = .51) and Job Seeking Skill tests (r = .51), with the Practical Reasoning: Zip Code and Paragraph Meaning tests falling just short of the .05 significance level.

In contrast to the above findings, an unusual trend in significant negative relationships results for two of the nonverbal cognitive skills tests. The finding might easily be written off as aberrant, by virtue of the small N's, if it did not persist as a similar pattern in both the Program-Completion and Post-Program (Not Employed) samples (two groups which, it should be noted, are conceptually more similar to one another than they are to a Post-Program group of enrollees who obtained full-time employment). For those two samples, Choose-A-Path scores are negatively related to the "Training Program Adjustment" scale at program completion (r = .54; Table 3) and to "General Community Adjustment" in the Post-Program period (for ex-enrollees Not Employed, r = -.39; Table 5). The Picture-Number test results in a similar pattern of negative correlations for those same criterion samples and scales--although barely short of significance in the case of the Program Completion sample (r = -.41; $p \cong .06$). Add to this on an overall pattern in the direction of negative validities for Technical and Outdoor interest scales with these same criteria and one tenuous clue to a sensible explanation of these results is possible. It is simply that an underlying sex differentiation was obscured by the necessity for combining male and female groups.

In essence, since both the Technical and Outdoor task preferences are known to be distinctly greater for the male enrollees, it is possible to assume that the remainder of this validity pattern is also sex dependent--i.e.,

that better performance for males on these two nonverbal cognitive tests is associated with poorer "Training Program Adjustment" (at the end of the program) and poorer "General Community Adjustment" (post-program) for those males who do not obtain employment. In partial support of this interpretation, it should be recalled that Choose-A-Path appeared as a masculine-oriented measure on the Masculine Vs. Feminine Interests factor.

One other point that should be noted in Tables 3, 4, and 5 is that, despite the relative independence of the Picture-Number test in relation to other cognitive skills, the measure tended to display significant (or near significant) predictive validities across the three criterion samples. Similar hints of the value of rote memory skill have appeared in its relation to job performance criteria with samples of Black Civil Service employees, as contrasted with lesser validity for White employees (Campbell, 1972).

Few of the remaining predictive validities are worth specific interpretation in such a small-scale study phase, except for mention of several that appear with the "Job Stability - Mobility Criterion" scale (Table 4). Among these ex-enrollees (Employed), higher scores on the Motivation for Vocational Achievement scale were associated with significantly greater mobility (in jobs, residence, and degree of job search effort; $r = .54$) along with a tendency--although not quite significant--for the more mobile enrollees to express lesser willingness to defer gratification ($r = .43$; $p < .10$). In addition, the more mobile the former enrollee the poorer his performance on the Arithmetic Skills test ($r = -.55$).

"Employment" as Criterion

In the above analyses it was necessary and logical to deal with different post-program criterion measures on the basis of whether former enrollees

did or did not obtain employment. This division, however, precluded the opportunity to examine the predictive value of the tests for what would normally be considered a major criterion measure applicable to any youth-work training program--i.e., Employment Vs. No Employment. By combining the two post-program samples into a single sample of 44 former enrollees it is possible to compute the biserial correlations between each test and the dichotomous employment criterion as shown in Table 6.

Insert Table 6 about here

None of the measures are found to be correlated significantly with employment as a criterion for this sample, nor are any trends discernible. The negative findings not only stand by themselves, they also buttress arguments against total dependence on a single, all-encompassing, program criterion of employment (as important as it may be). Limiting test validation to such a single outcome measure would have failed to point up clues to the predictive potential of the measures, as found here, when a wide variety of post-program performance capabilities are used to define end-of-program or post-program "success."

Conclusions

Measures of the Enrollee Test Battery, designed specifically for use with disadvantaged adolescents, appear fairly coherent and logical in their patterns of relationships with each other and with conventional cognitive skills tests. Substantial levels of concurrent validity for four of the subtests in the ETB--each of which require knowledge or reasoning capability--confirm their intended cognitive dimensional quality. At the same time, noncognitive attitude and interest scales yield factors that are

interpretable and relatively independent of cognitive skills. Demonstration of the relative uniqueness of such dimensions is, of course, desirable for measurement purposes (i.e., serving as a form of construct validity) and is believed most readily obtainable, for an educationally disadvantaged group with poor reading ability, when verbal skill (i.e., reading) requirements are removed as a primary component of test response.

The more critical questions of test value, however, can only be approached on the basis of their correlations with a broader range of performance criteria, and in a predictive framework. Within the limitations of very small and restricted samples, the results of these preliminary analyses have suggested a dominant predictive role for the various cognitive skills measures with several criterion dimensions. Special interest was focused on the rote memory measure because of its predictive possibilities in spite of its unique standing (i.e., low correlations) with the other cognitive skills tests. Unfortunately, a number of the cognitive test validities and their interpretations were believed to be confounded, in important ways, by having combined males and females in the same sample. The obvious need in future research is to obtain follow-up samples of sizes that are sufficient to allow for separate analyses by sex (recognizing the cost and difficulties in doing so when the respondents are minority group, school dropouts, residing in poverty level urban areas).

Patterns of usable predictive validities for attitude and interest scales might be inferred from the results (e.g., for Technical and Outdoor interest, Self-Esteem, Motivation for Vocational Achievement). But, for the most part, those findings should be considered inconclusive without larger, more representative, longitudinal samples and more definitive analyses by sex.

Part of any continued development of measures in the Enrollee Test Battery must also include refinement of the criterion scales (e.g., based on criterion relevance or equivalence), as well as the derivation of useful scoring profiles, the weighting of tests for optimum prediction of multiple criteria, and the development of normative data--all of which are to be incorporated in study efforts presently being planned or in progress.

References

Campbell, J. Personal communication, July 1972.

College Entrance Examination Board. Comparative guidance and placement program, Form RPG. New York: College Board, 1969.

Cronbach, L. J. Essentials of psychological testing. (2nd ed.) New York: Harper, 1960.

Eells, K. et al. Intelligence and cultural differences. Chicago: University of Chicago Press, 1951.

Flaugher, R. L. Minority versus majority group performance on an aptitude test battery. Research Bulletin 71-48. Princeton, N. J.: Educational Testing Service, 1971.

Flaugher, R. L., & Rock, D. A. Patterns of ability factors among four ethnic groups. Research Memorandum 72-7. Princeton, N. J.: Educational Testing Service, 1972.

Freeberg, N. E. Development of evaluation measures for use with neighborhood youth corps enrollees. Project Report 68-5. Princeton, N. J.: Educational Testing Service, 1968.

Freeberg, N. E. Assessment of disadvantaged adolescents: A different approach to research and evaluation measures. Journal of Educational Psychology, 1970, 61, 229-240.

Freeberg, N. E., & Reilly, R. R. Development of guidance measures for youth-work training program enrollees. Phase I: Measurement of program objectives and the development of criteria. Project Report 71-13. Princeton, N.J.: Educational Testing Service, 1971.

French, J., Ekstrom, R., & Price, L. A. Manual for kit of reference tests for cognitive factors. (June 1969 ed.) Princeton, N. J.: Educational Testing Service.

Harcourt, Brace, & World. The Stanford Achievement Test, Form W, Intermediate I. New York: Author, 1964.

Harcourt, Brace, Jovanovich. The Adult Basic Learning Exam. New York: Author, 1970.

Kaiser, H. F. The varimax criterion for analytic rotation in factor analysis. Psychometrika, 1958, 23, 187-200.

Karlsen, B., Madden, R., & Gardner, E. ABLE Handbook. New York: Harccurt, Brace, Jovanovich, 1971.

Karp, J. M., & Sigel, I. Psychoeducational appraisal of disadvantaged children. Review of Educational Research, 1965, 35(5), 401-412.

Knapp, R. R., Grant, B., & Demos, G. D. California occupational preference survey. San Diego: Educational and Industrial Testing Service, 1966.

Rohwer, W. D., Jr. Learning, race, and school success. Review of Educational Research, 1971, 41(3), 191-210.

Society for the Psychological Study of Social Issues (SPSSI). Guidelines for testing minority group children. SPSSI, 1964, 20(2), 127-145.

Tyler, L. E. The psychology of human differences. (2nd ed.) New York: Appleton-Century-Crofts, 1956.

Table 1

Intercorrelations of Test Scores for Youth-Work Training Program Enrollees (N ≈ 70[a])

	Job Knowledge (JK)	Practical Reasoning: Zip Coding (PRZ)	Practical Reasoning: Map Reading (PRM)	Job Seeking Skills (JSS)	Interest Scales: Clerical (CL)	Service (SE)	Technical (TE)	Outdoor (OU)	Science (SC)	Business (BU)	Aesthetic (AE)	Voc. Aspirations Minus Voc. Plans (VA-VP)	Attitude Towards Authority (AA)	Self-Esteem (SE)	Deferred Gratification (DG)	Motivation for Vocational Achievement (MVA)	Job Holding Skills (JHS)	Adult Basic Learning Exam: Vocabulary	Arithmetic	Stanford Achievement (Paragraph Meaning) (PM)	Choose-A-Path (CAP)	Figure Analogies (FA)	Picture Number (PN)
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1. JK		.31*	.25	.49	-.15	-.15	.03	.13	.14	.05	-.06	-.05	.22	.18	.09	.24	.44	.32	.37	.44	.29	.37	.16
2. PRZ			.51	.50	-.14	-.25	-.13	-.22	.12	-.04	-.16	-.02	.16	.09	.17	.32	.21	.41	.47	.61	.29	.54	.06
3. PRM				.45	-.05	-.07	.10	.07	.20	.16	.05	-.03	.30	-.04	-.05	.12	.12	.42	.45	.52	.38	.50	.05
4. JSS					-.13	-.16	-.35	-.32	-.08	-.12	-.19	.03	.25	.07	.16	.26	.31	.57	.58	.75	.32	.42	.24
5. CL						.16	.15	.04	.30	.47	.34	.10	-.01	.13	.09	-.01	.03	.05	-.05	-.02	-.06	-.09	-.06
6. SE							.10	.09	.22	.20	.45	.09	.06	.24	-.03	.29	.26	-.05	-.03	-.16	.02	-.19	.06
7. TE								.79	.58	.60	.54	-.14	.13	-.07	-.13	-.10	-.15	-.05	-.17	-.34	.19	-.10	.18
8. OU									.38	.41	.51	.11	.07	-.07	-.08	-.18	-.18	-.06	-.11	-.37	.33	.00	.07
9. SC										.55	.53	-.05	.16	.04	.03	.17	-.01	.01	.09	.09	.18	.17	.05
10. BU											.66	-.12	.05	.08	.08	-.02	.02	.14	-.04	-.04	.14	-.03	.10
11. AE												-.01	.13	.06	.03	.06	-.07	.04	-.01	-.16	.27	.00	.05
12. VPA													-.14	.09	.19	.05	-.01	.05	.25	.12	.15	.05	.04
13. AA														.26	.00	.17	.23	.12	.27	.31	.28	.19	.11
14. SE															.05	.42	.47	-.04	.26	.11	.21	-.15	.15
15. DG																.16	-.02	.08	-.01	.21	-.07	.08	-.01
16. MVA																	.22	.16	.15	.28	.29	.05	.25
17. JHS																		.17	.44	.34	.12	.12	.09
18. VOC																			.35	.54	.40	.46	.33
19. ARTH																				.60	.43	.43	.27
20. PM																					.21	.54	.19
21. CAP																						.38	.17
22. FA																							.14
23. PN																							--

[a] Range of N's = 67 to 74.

* Correlations of (approximately) .23 or greater are significant at the .05 level.

Table 2

Loadings on Varimax Factors for Total Test Battery[a]

	Tests	I Cognitive Ability	II Interest in Vocational Tasks	III Positive Social Attitudes	IV Masculine Vs. Feminine Interests
1.	Job Knowledge	53	-05	28	21
2.	Practical Reasoning Zip Coding	70	-09	07	-16
3.	Practical Reasoning Map Reading	68	14	02	03
4.	Job Seeking Skills	74	-22	-23	-19
5.	Voc. Interests: Clerical	-12	49	06	-32
6.	Voc. Interests: Service	-28	35	48	-06
7.	Voc. Interests: Technical	-05	71	-15	55
8.	Voc. Interests: Outdoor	-03	54	-15	69
9.	Voc. Interests: Science	16	71	05	05
10.	Voc. Interests: Business	05	82	00	-03
11.	Voc. Interests: Aesthetic	-05	80	10	10
12.	Voc. Aspirations Minus Voc. Plans	04	-02	12	-07
13.	Attitude Towards Authority	30	11	28	15
14.	Self-Esteem	-03	04	70	01
15.	Deferred Gratification	10	06	04	-30
16.	Motivation for Achievement	18	05	57	-11
17.	Job Holding Skills	22	-08	63	-03
18.	ABLE Vocabulary	65	08	06	-11
19.	ABLE Arithmetic	64	-07	37	02
20.	Stanford Achievement PM	79	-11	21	-37
21.	Choose-A-Path	50	19	25	37
22.	Figure Analogies	73	01	-10	-01
23.	Picture Number	24	07	24	15

[a]Decimal points omitted.

Table 3

Predictive Validities for 23 Tests with Three Program-Completion Criterion Scales

	Tests	Criterion Scales: Training Program Adjustment (N = 20)	Social Adjustment (N = 20)	Job Aspiration Level (N = 19)
1.	Job Knowledge	-.23	.19	.28
2.	Practical Reasoning Zip Coding	.21	.51*	-.02
3.	Practical Reasoning Map Reading	.34	.34	-.14
4.	Job Seeking Skills	-.23	.16	.02
5.	Voc. Interests: Clerical	.19	.28	-.20
6.	Voc. Interests: Service	.30	.04	-.19
7.	Voc. Interests: Technical	-.38	-.28	.14
8.	Voc. Interests: Outdoor	-.41 (p ≈ .06)	-.44	.08
9.	Voc. Interests: Science	.07	.14	.05
10.	Voc. Interests: Business	-.10	.01	.07
11.	Voc. Interests: Aesthetic	-.01	-.20	.02
12.	Voc. Aspirations Minus Voc. Plans	.22	.16	-.31
13.	Attitude Towards Authority	.16	.31	.10
14.	Self-Esteem	-.14	.47*	.28
15.	Deferred Gratification	-.02	-.11	-.09
16.	Motivation for Achievement	-.29	.12	-.14
17.	Job Holding Skills	.15	.25	.04
18.	ABLE Vocabulary	-.05	.09	-.17
19.	ABLE Arithmetic	.09	.66*	.26
20.	Stanford Achievement PM	.08	.52*	.05
21.	Choose-A-Path	-.54*	-.09	.01
22.	Figure Analogies	-.04	-.10	.14
23.	Picture Number	-.41 (p ≈ .06)	.07	.25

*Significant at .05 level.

Table 4

Predictive Validities for 23 Tests with Three Post-Program Criterion Scales (Employed Sample)

	Tests	Criterion Scales: Job Success & Adjustment (N = 15)	Striving for Vocational Success (N = 10)	Job Stability-Mobility (N = 15)
1.	Job Knowledge	.04	.04	.16
2.	Practical Reasoning Zip Coding	.46	-.31	-.07
3.	Practical Reasoning Map Reading	.18	-.57	-.12
4.	Job Seeking Skills	.51*	-.32	.18
5.	Voc. Interests: Clerical	.09	.36	-.28
6.	Voc. Interests: Service	-.27	-.02	.39
7.	Voc. Interests: Technical	-.25	.12	-.04
8.	Voc. Interests: Outdoor	-.36	.05	.06
9.	Voc. Interests: Science	.09	.21	-.02
10.	Voc. Interests: Business	-.39	.18	-.51
11.	Voc. Interests: Aesthetic	-.28	.06	-.12
12.	Voc. Aspirations Minus Voc. Plans	-.16	-.38	.10
13.	Attitude Towards Authority	.15	-.86*	.25
14.	Self-Esteem	-.30	-.05	.07
15.	Deferred Gratification	.26	-.38	-.43
16.	Motivation for Achievement	-.17	-.18	.54*
17.	Job Holding Skills	.26	.15	-.06
18.	ABLE Vocabulary	.49*[a]	-.26[b]	-.27[b]
19.	ABLE Arithmetic	.51*[a]	.11[b]	-.55*[a]
20.	Stanford Achievement PM	.43[a]	-.26[b]	-.12[a]
21.	Choose-A-Path	.01[a]	-.11[b]	-.17[a]
22.	Figure Analogies	.57*[a]	-.39[b]	-.25[a]
23.	Picture Number	.51*[a]	-.08[b]	-.18[a]

[a] N = 17; [b] N = 11.

*Significant at .05 level.

Table 5

Predictive Validities for 23 Tests with Three Post-Program Criterion Scales (Not-Employed Sample)

	Tests	Criterion Scales: General Social Adjustment (N = 27)	Striving for Personal Improvement (N = 17)	Realistic Aspiration (N = 27)
1.	Job Knowledge	-.03	.10	.01
2.	Practical Reasoning Zip Coding	.04[a]	-.12	.13[a]
3.	Practical Reasoning Map Reading	-.18[a]	-.34	.09[a]
4.	Job Seeking Skills	-.12[a]	.27	-.19[a]
5.	Voc. Interests: Clerical	.09	.24	.10
6.	Voc. Interests: Service	-.31	-.19	-.31
7.	Voc. Interests: Technical	-.45*	-.12	-.05
8.	Voc. Interests: Outdoor	-.36	-.12	-.05
9.	Voc. Interests: Science	-.26	-.24	-.05
10.	Voc. Interests: Business	-.25	.17	-.06
11.	Voc. Interests: Aesthetic	-.29	.22	-.03
12.	Voc. Aspirations Minus Voc. Plans	.00	-.38	.16
13.	Attitude Towards Authority	-.19	-.42	-.29
14.	Self-Esteem	-.06	-.17	-.32
15.	Deferred Gratification	.09	-.22	-.03
16.	Motivation for Achievement	.00	-.30	.14
17.	Job Holding Skills	-.11[a]	-.13	-.05[a]
18.	ABLE Vocabulary	-.20[b]	.03	-.04[b]
19.	ABLE Arithmetic	-.16[b]	-.09	-.10[b]
20.	Stanford Achievement PM	.12[b]	-.13[d]	.28[b]
21.	Choose-A-Path	-.39*[c]	-.08[d]	-.24[c]
22.	Figure Analogies	.06[c]	-.48*[d]	.17[c]
23.	Picture Number	-.39*[c]	.02[d]	-.06[c]

[a]N = 26; [b]N = 25; [c]N =24; [d]N = 15.

*Significant at .05 level.

Table 6

Predictive Validities of 23 Tests Using Employment Vs. No Employment as Criterion

(N = 44)

	Test		r_{bis}
1.	Job Knowledge		.09
2.	Practical Reasoning Zip Coding		-.15
3.	Practical Reasoning Map Reading		.19
4.	Job Seeking Skills		.11
5.	Voc. Interests	Clerical	.26
6.		Service	.04
7.		Technical	.08
8.		Outdoor	.22
9.		Science	.05
10.		Business	.14
11.		Aesthetic	.11
12.	Voc. Aspirations Minus Voc. Plans		.00

	Test	r_{bis}
13.	Attitude Towards Authority	.26
14.	Self-Esteem	.12
15.	Deferred Gratification	.10
16.	Motivation for Achievement	-.25
17.	Job Holding Skills	-.07
18.	ABLE - Vocabulary	.16
19.	ABLE - Arithmetic	.01
20.	Stanford Achievement	.19
21.	Choose-A-Path	.28
22.	Figure Analogies	.14
23.	Picture Number	-.22